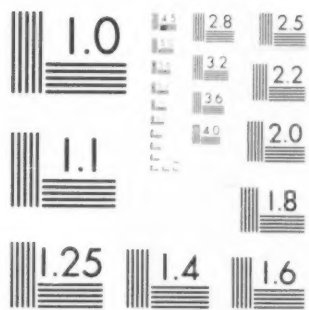


2



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DOCUMENT RESUME

ED 152 291

IB 005 772

AUTHOR Hunka, S.; And Others
TITLE Report on the Use of the Computer-Assisted Instruction Course STAT1 as Used in Educational Psychology 502. 1975-1976.
INSTITUTION Alberta Univ., Edmonton. Div. of Educational Research Services.
REPORT NO DERS-01-102; RIR-76-7
PUB DATE Nov 76
NOTE 26p.
EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.
DESCRIPTORS *Computer Assisted Instruction; Course Descriptions; *Course Evaluation; Higher Education; *Instructional Design; *Research Methodology; Tables (Data)

ABSTRACT

STAT1 is a computer-assisted instruction course in basic research methodology in education. Student performance data were collected from testing, revision, and retesting. This study includes improvements to STAT1, contents of STAT1 as used in the summer of 1976, a description of examinations and laboratory assignments, characteristics of students, a comparison of marks for the 1975 and 1976 groups, structural changes among the relationships of examinations from 1975 to 1976, attitude of students towards STAT1, general observations, and current use of STAT1. (VT)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED152291

November 1976

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

RIR-76-7

REPORT ON THE USE OF THE
COMPUTER-ASSISTED INSTRUCTION COURSE
STAT1
AS USED IN EDUCATIONAL PSYCHOLOGY 502
1975-1976

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

E.W. Romaniuk

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

S. Hunka
E.W. Romaniuk
T.O. Maguire

Division of Educational Research Services
Faculty of Education, The University of Alberta
Edmonton, Alberta T6G 2G5

IR005-772

CONTENTS

	<u>Page</u>
Introduction	1
Improvements to STAT1	1
Contents of STAT1 as Used in the Summer of 1976	4
Examinations and Laboratory Assignments in STAT1	7
Characteristics of Students in the 1976 Testing	8
Comparison of Marks for the 1975 and 1976 Groups	10
Structural Changes Among the Relationships of Examinations from 1975 to 1976	12
Attitude of Students Towards STAT1.	15
General Observations	18
Current Use of STAT1	21
APPENDIX A	22
APPENDIX B	23

Introduction

STAT1 is a computer-assisted instruction course designed by the authors to provide instruction in basic research methodology required of students intending to carryout research in education. The content of the course is taught in the context of research problems typically found in education.

The organization of the content is quite linear in that sections of the course are internally linked together within the computer code. Provisions are provided to allow a student to bypass the internal order of the content by use of special keyboard operations which allow access to a new, or a previously covered section of the course. In many sections of the course the student is formally presented with the option of reviewing important concepts. This facility was not provided during an examination. Extensive answer analysis hierarchies are built into the course for diagnostic purposes. The current version of the course contains about 100,000 computer instructions.

The first testing of the course was made during Summer Session of 1975, while a second testing of the course, after improvements, was made in the Summer Session of 1976. Approximately 100 students took various portions of STAT1 during the winter session of 1976 under the guidance of instructors other than the authors. No student performance data were collected for this latter group. The period between September 1975 and June 1976 was used for purposes of improving STAT1 based upon the student performance data obtained during the summer of 1975.

The authors have crudely estimated that it has taken 2,500 man-hours to bring the course to its present state of development. An estimate of the development costs of this course, not counting academic time used for detailing the course content and course code, is \$10,000. The majority of academic time spent on the development of this course was outside of regular working hours.

Improvements Made to STAT1

Based upon the records of student performance collected during the summer of 1975, STAT1 was improved. Improvements were based upon the analysis of over 100,000 records detailing every answer each student made to questions of the course.* The records when aggregated for each question provided the following information printed in a form easily understandable by an author:

- a) -the location in the course of the question being asked
- b) -the specific question being asked, including a re-creation of the complete screen display before the student

*The average number of responses made in the course by each student was about 3,000. Any critics who claim the course is not sufficiently interactive should keep this figure in mind.

- c) -the answers which were to be identified as acceptable and unacceptable. (These may have included specific and precise combination of words, number limits, or skeleton structures of phrases.)
- d) -the frequency of student responses which fell into each identifiable answer category as detailed in (c) above for the first, second, third, fourth, etc., attempts.
- e) -the frequency of the response which could not be identified as well as a list of such responses.
- f) -the I.D. number of each student, and the response given identified by the order in which they were given.

The analysis detailed above was of necessity run on the larger computing system located in Computing Services. The procedure for obtaining the aggregation of the student data records involved having the computer re-create the contents of the screen display and answer analysis procedures used in STAT1, by scanning through the Coursewriter source code in which STAT1 was written. The information produced by this first step was then merged with the sorted records of student performance data and analyzed. The sorting provided the collation of all responses and their time base for each question.

The information obtained by the aggregation of student performance data provided the basis for improving the course in the following manner:

- identification of responses to which the computer did not provide an adequate response or feedback message
- identification of poorly worded questions and text
- an indication of the location in the course where instructional improvements were required
- identification of student answers which were indicative of unanticipated misconceptions created by the course presentation

In addition to the analysis of student performance data, all the comments students left for the authors were studied to identify further, those portions of the course which required improvement, but which could not be identified by the analysis of student performance data. For example, a comment left by a student such as, "I think you should have given statistical tests separately from confidence intervals," identifies a problem in the ordering of the subject matter which could not have been as easily identified by analysis of responses to specific content questions. Other improvements were carried out which involved redesigning the answer analysis hierarchy which followed each question. Since the domain of possible answers of the student can input is almost infinite, while the answers which can be identified, are finite, the analysis of the student's input must be handled in a manner which will direct him to converge upon the correct answer. In addition, wrong answers which can be identified as evidence of a particular type of error, must also be corrected within the same process.

On the basis of the observations made during the summer of 1975, Mr. N. McGinnis, system analyst, provided the CAI system with a 'desk calculator' facility which could be used by the student when, in the context of instruction, numerical calculations were required. This facility obviated the need for students to purchase a pocket calculator or to borrow one from the Test Library operated by the Department of Educational Psychology.

The content of STAT1 was changed during the winter of 1975 by the addition of two new chapters. One chapter provides more detailed instruction and practice in calculating areas under the normal curve, and in calculating probabilities associated with the binomial distribution. On the basis of observations made during the summer of 1975, it was concluded that students had too many difficulties in reading the normal curve table in their texts. The other new chapter was designed to provide the student with instruction in the following content areas:

- the need for precision in specifying how data was to be analyzed
- the primary steps in an experimental design
- the importance of different scales of measurement to the method of analysis
- encoding the information of a questionnaire for entry onto a computer data card
- how the computer at Computing Services handles data submitted to it under "time-sharing"
- selecting a computer program from the DERS program library
- how to construct computer format statements using Fortran which describes to the computer the data to be read from the data card
- how to run a computer program using the Computing Services Amdahl 470/46 computer which operates under MTS

In addition to the 'attitude-free' content described above, Dr. Hunka has clearly indicated his own attitude towards those researchers who make important research decisions based on computer output, even though they cannot be sure that the computer analyzed the correct data on the data card.

Further improvements were made to STAT1 by providing the student with a complete list of 'restart' points in the course. With this information the student was not only able to locate his position in the course, he could also move himself about the course for purposes of review. In addition to this information, the students were provided with the minimum, average, and maximum times students from the previous summer took to complete each section of the course.

In the original testing of STAT1, the course directed the student to complete seven laboratory assignments which paralleled the content of the CAI course. During the winter of 1975-76 the seven assignments were rewritten to improve their content and format, and three further assignments were added.

Any examination administered under the CAI mode raises the question of test security. This concern was evident to the authors of STAT1 since the final testing. Students were able to copy down questions from the screen and to then record whether their answers were right or wrong. In the 1976 testing of the course, questions were first randomized before presentation to the student. Thus, although each student received the same questions, the order of presentation was different for each. It is hoped that further revisions will provide an examination based upon randomization without replacement from a test item pool. The provision of immediately scoring each item as right or wrong was retained in the 1976 version, as was the immediate summary of the examination results which identified areas of the student's weakness. The pencil and paper examinations were also improved during the winter of 1975-76, and a new examination on research design was developed. As students completed the CAI portion of the course, a seminar was given on the topic of research design. This was required because the material was not fully available in the computer version of the course. The reference text used was Stanley and Campbell's Experimental and Quasi-Experimental Designs for Research.

In order to better assess the student reaction to each major section of the CAI course, Likert-type scales were introduced prior to each examination presented by the computer, and as each major topic was finished by the student. This was done to provide more objective evidence of the reaction of students to the course than that provided by some instructors who only used the course on a casual basis during the winter 1975-76.

Contents of STAT1 as Used in the Summer of 1976

Table I details the contents of STAT1 as it was used in the Summer Session of 1976. Of necessity, detailed descriptions cannot be given here for a course as large as STAT1.

Examination of Table I, and Table III indicates that the course is not covering trivial content. The authors realize that in making the course a CAI course, they are opening the door to what they think should be taught, as well as what they feel are acceptable methods of instruction. During the past one and one-half years, this position has led to many healthy discussions with other staff members.

TABLE I
CONTENTS OF STAT1 AS OF JULY 1976

Segment Number	Title	Notes
0	Introduction to Terminal	shows the student how to use the terminal including sign-on and off procedures
3	Descriptive	data tabulation, measures of central tendency, variability, parameters of distributions, transformation of data, weighted means, deviation and Z scores
4	Prediction	equation of straight line, least square error estimators of parameters, calculus option, scatter diagram, making predictions
4	Correlation	relationship to regression, covariance, effects of linear transformations, proportion of variance predictable calculations
4	Correlation	Biserial, Point Biserial, Phi, and relationship to Pearson correlation
20	Normal Curve	calculations of areas for any portion of the normal curve
20	Binomial Distribution	combinations, calculating probabilities, applications, Pascal Triangle
5	Multiple and Partial Correlations	relationship to ordinary prediction, model, calculus option, matrix equations inverse matrix, solution of weights by two methods, simulation problem, partial correlation as correlation of parts, cautions on use
6	Statistical Tests and Confidence Intervals	statistical tests for means, mean differences, correlations, proportions; use Z, t, Fisher's Z. Confidence intervals for mean differences, correlation, proportion
7	Terms of Inferential Statistics	population, sample, parameter, statistic application
8	Random Sampling Distribution of the Mean	effects of increasing sample size on accuracy of estimating a parameter, RSD of mean, standard error of mean

TABLE I (Cont'd)
 CONTENTS OF STATI AS OF JULY 1976

Segment Number	Title	Notes
9	Probabilities of Sample Means	various examples of differing sample size on distribution of means, limits
10	Hypothesis Testing	type I and II errors, decision rules, null and alternate hypothesis, power curve, effects of sample size
11	Anova 1-way Purpose	purpose, examples, factors, dependent and independent variables
12	Anova 1-way Computation	estimates of variance components, degree of freedom, mean square
13	Anova 1-way Intuitive Approach	RSD of mean, example of three samples drawn from three populations, independence of variance estimates, between and within group variance, F
14	Anova 1-way Mathematical	algebraic model and parameters, treatment effect, error, mean effect
16	Tukey and Scheffe	rationale, relation to 't' tests, use of studentized 't' table, examples, contrast coefficients, Scheffé statistics, comparison of both methods
17	2-way Interaction in Anova; Interpretation	additivity concept, ordinal and disordinal interaction, plotting effects
19	Chi-Square	rationale, expected and observed frequencies, degree of freedom, contingency tables, Yates correction
25	Data Analysis and Research Design	planning steps, design of experiment, scales selecting valid parameters, encoding data into a data card, computer format statements, use of DERS programs, running a job under MTS

Examinations and Laboratory Assignments in STAT1

The current version of STAT1 requires the student to complete eleven examinations and ten laboratory assignments. The nature of the examinations are detailed in Table II.

TABLE II
EXAMINATIONS IN STAT1

Topic	C/R*	Number of Items
Descriptive Statistics	C	57
Prediction	C	33
Correlation #1	C	31
Correlation #2	C	23
Hypothesis Testing	C	22
Anova	R	16
Tukey and Scheffé	R	15
Interaction	R	14
Chi-Square	R	6
Data Analysis and Research Design	R	20
Final Examination	R	70

*C refers to a computer administered test.

R refers to a regular pencil-paper administered test. In this category some were multiple choice, while others were of the problem solving (computational) type.

Laboratory assignments in which calculations had to be carried out for a specific problem or problems, were given in the following areas:

1. Descriptive Statistics
2. Calculating areas under the Normal Curve and the Binomial Distribution
3. Prediction and Correlation
4. Hypothesis Testing
5. Statistical tests
6. Analysis of Variance 1-way
7. Tukey and Scheffé
8. Analysis of Variance 2-way
9. Chi-Square
10. Data Analysis

All the laboratory assignments were checked by an instructor of the course. The CAI program indicated to each student when an assignment or examination could be taken. Thus, this work was completed at different times in the course depending upon the speed at which the student was able to progress through the material.

Characteristics of Students in the 1976 Testing

Of the thirty-seven students for whom data was maintained by the computer, thirty-four were taking STAT1 as a formal credit course identified in the University Calendar as Educational Psychology 502. Three students were taking the course for purposes of their own interest and review. Twenty-two students were registered in the Department of Educational Psychology, three in Elementary Education, four in Secondary Education, three in Industrial and Vocational Education, one in Educational Foundations, and four were from departments outside of the Faculty of Education. Twenty-seven of the students had not taken a statistics course in the last five years, while only twelve had taken at least one mathematics course since leaving high school. Two of the thirty-seven students had taken a course by CAI before taking STAT1.

To assess the attitude of each student towards working with numbers, each student was asked to rate his reaction to the question:

"I have no difficulty in working with numbers
and concepts involving numbers."

on a seven point scale. The distribution of the responses was as follows:

	Strongly Disagree				Strongly Agree		
	1	2	3	4	5	6	7
Frequency:-	4	7	4	7	8	5	2

Comparable data for the 1976 group was not collected. However, some self selection of the 1976 group could have occurred since this group was fully aware that the course would be offered in a CAI format.

Comparisons of Results for the 1975 and 1976 Groups

Time to Complete Sections of the Course

Table III details the average, minimum, and maximum time students took to complete various chapters in the course. The time to complete each chapter includes any time which the student may have used to cover and review the material.

For the 1975 testing, the examinations for the Descriptive and Prediction and Correlation sections were contained within the tutorial portions of the course, while in 1976 the examinations were separated as independent units of the course. Thus, to make the 1975 data comparable to the 1976 data for these sections, the examination times must be added to the instruction times for these sections, e.g., the average time for the Descriptive section is 3.2 hours (2.1+1.1), and 12.9 (9.9+1.0+1.1+.9) for the Prediction and Correlation section.

TABLE III
TIME TO COMPLETE VARIOUS SECTIONS OF STATI
BY STUDENTS IN 1975 and 1976

Topic	Average Hours		Minimum Hours		Maximum Hours	
	1975	1976	1975	1976	1975	1976
Introduction to Terminal	-	.6	-	.2	-	1.5
Descriptive Statistics	6.0	2.1	2.7	1.0	12.0	5.9
Prediction & Correlation	12.0	9.9	6.5	5.3	19.0	20.8
Multiple & Partial Correlation	2.9	3.3	.8	2.1	4.0	5.0
Tests & Confidence Intervals	2.0*	4.8	1.0	2.5	6.0	8.5
Terms of Inferential Statistics	.1	.3	.1	.1	.6	.7
Random Sample of Means	1.1	1.1	.5	.3	2.4	2.3
Probability of RSD of Mean	.9	1.0	.2	.3	2.6	6.1
Hypothesis Testing	7.7	6.1	2.3	2.6	15.2	17.5
Anova: Purpose	1.1	1.0	.4	.3	2.0	1.6
Anova: Computation	2.0	1.4	.7	.6	4.4	2.8
Anova: Intuitive Approach	2.5	2.2	1.1	.7	4.1	4.1
Anova: Mathematical	1.2	1.3	.3	.7	4.4	2.0
Tukey & Scheffé	6.6	6.3	2.4	3.5	10.2	11.2
Anova: 2-way Interaction	3.5	3.6	1.3	1.3	5.7	6.1
Chi-Square	1.3	1.6	.5	.8	2.4	2.3
Normal Curve and Binomial	-**	5.7	-	3.0	-	11.6
Data Analysis & Research Design	-**	3.1	-	1.0	-	5.8
Examinations						
Descriptive Statistics	-	1.1	-	.4	-	3.2
Prediction	-	1.0	-	.3	-	2.2
Correlation #1	-	1.1	-	.4	-	2.8
Correlation #2	-	.9	-	.3	-	2.4
Hypothesis Testing	-	.9	-	.4	-	3.0
TOTAL HOURS	50.9	60.4	20.8	8.1	95.0	129.4

*In 1975 the topic of confidence intervals was not part of this section.

**These two sections were added to the 1976 testing and did not appear in the 1975 version of the course.

On the basis of the current content of STAT1, the 1976 data indicate that the average time to complete the course is 60.4 terminal-hours, with a minimum of 28.1 terminal-hours, and a maximum of 129.4 terminal-hours. This course is listed in the University Calendar as a three hour lecture and three hour laboratory course, and on the basis of a seven month year, this would amount to about 84 hours of lectures and an additional 84 hours of laboratory work. Thus, on the basis of the data of Table III, it would appear that the instructional component of STAT1 which requires 60.4 terminal-hours, represents a saving of about 24 hours of an academic staff member's time. In addition, however, since the instructors in both 1975 and 1976 handled all the laboratory assignments during the time the course was operational on the computer, an additional 84 hours of laboratory work was not required on the part of the instructor. Since the instructor assisted students at terminals as required, time was available to mark and discuss laboratory assignments with each student on an individual basis. Although this analysis and comparison of times may be open to debate, since an instructor of the regular method might argue that much more material is being covered by the regular lecture approach, the authors feel that the burden of proof must be placed on the lecture method. This position is taken because the content as well as the method of instruction contained within STAT1 is precisely detailed and open for inspection.

In June 1976, three students were allowed to start the course before the regular Summer Session began. Although this may appear administratively improper, it did serve the very valuable function of having the corrections which were made to STAT1 in the winter of 1975-76, tested before the students arrived in July 1976. These three students were able to complete the course within the month of June by taking full advantage of available terminal time. These students were required to attend the seminars in research design during the same period of time as the regular students, and to take the final written examination with other students during the last week of Summer Session.

Certainly, the maximum time of 129.4 terminal-hours to complete the course is considerable, and the authors will speculate more about the reasons for this later in the paper.

Comparisons of Marks for the 1975 and 1976 Groups

A cursory study of the means and variances of the examination marks as given in Table IV indicates that the marks on the Descriptive and Prediction examinations are lower in 1976 than in 1975, while those of Correlation #1, Correlation #2, Hypothesis Testing, and Chi-Square, are higher in 1976 than in 1975. Valid comparisons of examination results can only be made for these examinations because modifications were made to test items of the other examinations. The variances of the marks in Descriptive Statistics is slightly higher in 1976 than in 1975, while in Prediction, Correlation #1, and Correlation #2 they are smaller in 1976 than in 1975. A considerable reduction in the variability of marks in Correlation #1 was observed in 1976. Increases in variability can be noted in Hypothesis Testing (8.72 to 9.96) and Chi-Square (7.87 to 9.59).

TABLE IV
STAT1 1975 and 1976
MEAN PERCENTAGES AND STANDARD DEVIATIONS

	Items	1975 (N=32)		Items	1976 (N=32)		NOTE
		Mean	Standard Deviation		Mean	Standard Deviation	
1. Descriptive	57	87.83	6.71	57	85.81	7.94	1
2. Prediction	33	84.75	7.71	33	83.20	7.54	1
3. Correlation #1	31	81.45	11.23	31	85.81	7.32	1
4. Correlation #2	23	84.51	9.41	23	87.95	8.12	1
5. Hypothesis Testing	22	88.50	8.72	22	90.10	9.96	1
6. Anova	18	76.22	11.64	16	79.46	10.83	2
7. Scheffe	15	81.04	13.55	15	77.33	16.19	2
8. Interaction	15	74.58	13.89	14	74.49	18.01	2
9. Chi-Square	6	92.00	7.87	6	93.90	9.59	1
10. Data Analysis and Research Design	15	64.79	12.61	20	73.29	12.10	3
11. Final Exam	46	74.91	10.25	70	72.45	10.58	4
12. Final Mark		80.06	5.95		81.00	7.4	

NOTES:

1. Examinations in these topics were identical in years 1975-76.
2. Examinations in these topics contained essentially the same items, but revised before use in 1976.
3. About 50% of the items were common between the two years.
4. Basically the same initial 46 items of 1975 were used, with an additional 24 items added to reflect a better proportion of topics in the course, plus items to cover new course material in the use of the normal curve, binomial distribution, and using the computer for data analysis.

Tests 1 through 5 were computer administered. All others were multiple choice pencil and paper type with the exception of Chi-Square which involved mainly calculations.

In order to determine whether the differences of means for the 1975 and 1976 groups were significantly different, the marks for the topics of Descriptive Statistics, Prediction, Correlation #1, Correlation #2, Hypothesis Testing, and Chi-Square were analyzed using Hotelling's T-squared statistic (a multivariate analog of the familiar 't' test.) The results indicated that a significant difference ($T\text{-squared}=15.137$, $df=(6,60)$, $p\leq .04$) in favor of the 1976 group, existed between the vectors of test means. A further analysis using 't' tests on each individual examination indicated that the examinations on Correlation #1 and Correlation #2 contributed most to significance of the differences. In addition a significant reduction was found in the variability on the Correlation #2 examination between 1975 and 1976 groups ($F=2.358$ $df=(31,34)$, $p\leq .02$).

Given that considerable time was spent improving the course prior to the 1976 testing, the fact that increased performance contributed most to the significance of the differences, is encouraging to the authors. The fact that a slight decrease in performance was noted for the examination marks in Descriptive Statistics and Prediction, although not a statistically significant decrease, is open to speculation. At the moment, the authors who handled the 1976 use of the course would offer their opinion that the decreased performance on the first two examinations was due to poor understanding of some rather basic arithmetic on the part of a number of students. As soon as these deficiencies were overcome by these students, in some cases with frustrations, their progress improved. The authors speculate further on this point at the end of the paper.

Structural Changes Among the Relationships of Examinations from 1975 to 1976

In order to get some idea of how improvements in the course, which were made before its use in 1976, affected the course as it was presented to the students, a factor analysis of the examination marks for both years was carried out. Since the factor analysis will reflect what the various tests measure as well as the extent to which students possess various competencies, and since the tests remained basically the same over the two testings as far as content was concerned, the changes in the factors obtained for the 1975 and 1976 periods should be due primarily to differences in the relationships among student competencies. Of course, differences in the samples of students would confound the results of interpretation of the factors.

Both the 1975 and 1976 results were factor analyzed using communality estimates because of the small number of variables. For both sets of data three factors were extracted as being representative of the most meaningful solution. Other number of factors were also tried, but these were discarded because factors beyond three were found to have such small loadings as to prevent a meaningful interpretation. The factors for both testings appear in Table V, VI. It should be noted that this analysis is independent of any mean differences which were noted earlier. Table V contains the test clusters formed by transforming the factors of 1975 to a configuration as close as possible to that of the 1976 factors.

TABLE V
FACTOR ANALYSIS OF 1975
EXAMINATION RESULTS

Test	Factor			h ²
	I	II	III	
Descriptive	.782	.057	.042	.616
Prediction	.224	.701	-.143	.561
Correlation #1	.288	.767	.004	.671
Correlation #2	.427	.446	-.225	.432
Hypothesis Testing	.711	.226	.220	.605
Analysis of Variance	-.063	.604	-.036	.370
Tukey and Scheffé	.424	.347	-.485	.536
Interaction	.556	-.022	-.371	.447
Chi-Square	.367	.087	-.138	.161
Research Design	.061	-.034	.416	.178
Final Examination	.673	.423	.154	.655

TABLE VII
FACTOR ANALYSIS OF 1976
EXAMINATION RESULTS

Test	Factor			h ²
	I	II	III	
Descriptive	.763	.336	.082	.702
Prediction	.280	.644	-.078	.499
Correlation #1	.705	.176	.236	.583
Correlation #2	.720	.383	.260	.732
Hypothesis Testing	.805	.096	.142	.678
Analysis of Variance	.132	.547	.337	.430
Tukey and Scheffé	.220	.761	.257	.693
Interaction	.424	.348	.569	.625
Chi-Square	.007	.084	.302	.099
Research Design	.404	-.129	.474	.405
Final Examination	.782	.425	.039	.794

TABLE VII
COMPARISON OF 1975 FACTORS ROTATED TO APPROXIMATE
THE 1976 FACTORING OF TESTS

1975 CLUSTER	FACTOR	1976 CLUSTER
Descriptive Statistics Correlation #2 Hypothesis Testing Final Examination —	I	Descriptive Statistics Correlation #1 Correlation #2 Hypothesis Testing Interaction Research Design Final Examination —
Prediction Correlation #1 Correlation #2 Analysis of Variance Tukey and Scheffé —	II	Prediction Analysis of Variance Tukey and Scheffe Final Examination —
Descriptive Statistics Tukey and Scheffé Interaction	III	Interaction Research Design

On the basis of Table VII it is more difficult to interpret the clustering of tests forming the factors for the 1975 results. For example, from a content point of view it is difficult to justify the communality between the test on interaction and Tukey and Scheffé. Correlation #1, although linked to Prediction, basically requires different skills which are not based directly on the linear model. Correlation #1 does require understandings of variance and covariance which are required for the calculation of the parameters in prediction. Correlation #2 covered other types of correlation, some of which can be derived from the basic Pearson model used in Correlation #1, however, the Correlation #2 test did not require these skills, concentrating mainly on the situations in which dichotomous and continuous measures could be used together to obtain a correlation coefficient, e.g., biserial. Thus, it was concluded that the factors defined by the clustering of the tests could not be readily supported by the curriculum design of the course.

Table VII also gives the test clusters for the 1976 sample, that is after improvements were made to the course. It is argued that this analysis is more acceptable. In the first factor can be found those examinations which require elementary type arithmetic computations, basic understandings of operators used in statistics, measures of central tendency and variability. The content of Research Design and Interaction examinations

required basic understanding of central tendency and variability. The Research Design test involved, in addition, the logic of experimentation, as well as common problems associated with data analysis. The Interaction test required understanding of the use of group means, plotting of these means, and interpretation of interaction effects. Thus, it had components which were both common with other tests loading on Factor I as well as components not in common. The ability to plot the means in the Interaction section of the course provides a common element with the plotting skills which are required under Prediction, and there is a small loading which may indicate this communality.

Since improvements were made to the content and operation of STAT1 prior to the 1976 testing, in most cases the examinations were identical (and where they were not because of the change of a few items, the skills being measured were certainly not changed), the authors would see the results of the factor analysis as reflecting a truer picture of the interrelationships among the content taught than that of the 1975 analysis. The improved relationship among the content taught in the 1976 testing is considered to be, to a large extent, due to the improvements made in the course.

A confounding variable in this analysis is represented by the fact that different authors wrote different sections of the course. However, the clustering of tests in either the 1975 or 1976 data cannot be clearly attributed to course authorship.

Attitude of Students Towards STAT1

Some instructors not directly involved with the preparation of the course but whose students had made use of portions of the course, indicated that some students reported deficiencies in the course, like for example, the course was overly boring. In the 1976 version of STAT1, data were collected to confirm or disconfirm the belief of the authors that such attitudes were not consistent with their observations. The authors recognize the possible interaction between the instructor's support to students and the degree to which the student may rate the course.

At the end of each segment, and always before an examination was given by the computer, the student was asked to rate the following statements on a seven point Likert scale:

- 1) "I found the material in this chapter was quite useful in helping me learn basic statistics."
- 2) "The material in this chapter was quite boring."
- 3) "I found this chapter easy to understand."
- 4) "I previously had not been taught any of the material in this chapter."
- 5) "There was never anyone available in the terminal room to help answer my questions about this chapter."

- 6) "I feel that someone who is knowledgeable in the course content of this chapter should always be available in the terminal room."
- 7) "Whenever there was a course content specialist available in the terminal room that person very often did not know the material very well."
- 8) "I feel that it would have been easier for me to learn the material in this chapter by classroom instruction rather than by computer-assisted instruction."

In order to bring the values of each scale to the same meaning, the questions numbered 2, 5, 7 and 8 were reflected. The means for each question as rated by the 1976 students are given in Table VIII. A rating towards the maximum value of 7 indicates a positive feeling towards the attitude being rated. Because questions 2, 5, 7 and 8 were reflected, the attitudes in the table have been reversed, e.g., question #2 is noted as measuring interest rather than "boringness."

Table VIII indicates that the chapters on Descriptive Statistics and Inferential Terminology were those which students indicated contained the least amount of new material. The chapters on Correlation #2, Multiple and Partial Correlation, Anova-mathematical, Anova-intuitive, and Interaction were viewed as containing the most amount of new material. The greatest variability of ratings existed for the question of whether the material had been encountered before. The chapters seen as being the most useful were Anova-purpose, Chi-Square, and Data Analysis section which also contained some elementary research design. The least useful chapters as seen by the students were Anova-intuitive and Anova-mathematical.

The most interesting sections of the course were Chi-Square, the purpose of Analysis of Variance, and the Data Analysis sections, while the least interesting were multiple and partial correlation, the intuitive and mathematical approach of analysis of variance.

The more difficult sections of the course as perceived by the students were the sections on Anova-mathematical, Anova-intuitive, and Statistical Tests and Confidence Intervals. The easiest were the sections on Inferential Terminology, Anova-purpose, and Chi-Square.

The question of availability of assistance to students while they were in the terminal room, was consistently rated high. No average rating fell below 6.00. This, it is hoped, reflects the authors' position that an instructor should let the computer take over the bulk of the instruction and step in to assist students on an individual basis. A similar consistently high rating was given to the question concerning the extent to which the person in the terminal room was knowledgeable in the course content. Variability on this question was also consistently small.

The question on the actual need for someone to be in the terminal room received average ratings between 5.2 and 6.36. the latter rating being for the section on Descriptive Statistics. The authors had hoped that ratings for this question would have been lower indicating that the course could "stand alone." It appears that this is not how the students

TABLE VIII
MEANS AND STANDARD DEVIATION OF RATINGS ON ATTITUDES TOWARDS STAT1: 1976

Topic	USEFUL		INTEREST		EASY		NEW		HELP AVAIL.		HELP ND.		HELP KNW.		CLASS-CAI	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Descriptive	5.94	.95	5.78	1.31	3.72	1.45	4.75	2.52	6.57	1.29	6.36	1.27	6.89	.31	5.66	1.47
Normal Curve & Binomial	5.94	1.11	5.52	1.58	3.86	1.50	5.57	2.07	6.85	.56	5.76	1.68	6.60	1.18	6.21	1.17
Prediction	5.94	.98	5.76	1.15	3.65	1.49	5.78	1.80	6.65	.76	5.97	1.61	6.89	.32	6.06	1.29
Correlation #1	5.35	1.35	5.29	1.56	3.51	1.66	6.00	1.43	6.68	.79	5.91	1.81	6.91	.28	5.62	1.51
Correlation #2	5.94	1.11	5.37	1.42	4.09	1.56	6.34	1.12	6.74	.69	5.80	1.75	6.85	.70	5.88	1.35
Multiple & Partial	5.25	1.20	4.77	1.62	3.06	1.49	6.63	.76	6.29	1.45	5.74	1.73	6.91	.29	5.60	1.52
Hypothesis Testing	5.33	1.46	5.50	1.03	3.09	1.68	6.05	1.69	6.18	1.61	6.27	1.32	6.77	.67	5.32	1.72
Stat. Tests & C.I.	5.56	1.26	4.83	1.58	2.83	1.78	5.53	2.01	6.39	1.34	5.86	1.73	6.89	.31	5.42	1.62
Inferential Terminology	6.23	1.26	5.76	1.52	6.76	.64	4.00	2.44	6.64	.73	4.74	2.39	6.74	.66	6.15	1.22
RSD of Mean	6.00	1.25	5.63	1.44	5.26	1.44	5.33	2.10	6.56	1.09	5.20	2.01	6.73	.66	5.88	1.20
Probability RSD of Mean	5.91	1.13	5.54	1.23	5.06	1.37	5.26	2.09	6.32	1.49	5.26	2.03	6.60	.73	5.89	1.28
ANOVA 1-way purpose	6.36	.88	6.48	.87	5.80	1.24	5.26	2.05	6.03	1.83	5.34	1.80	6.50	.95	5.26	.95
ANOVA Computation	5.09	1.61	4.94	1.57	3.15	1.79	6.03	1.82	6.44	1.42	5.71	1.84	6.74	.66	5.50	1.44
ANOVA Intuitive	3.88	1.53	4.77	1.44	2.40	1.34	6.53	1.17	6.09	1.65	5.91	1.73	6.88	.53	4.94	1.80
ANOVA Mathematics	4.13	1.54	4.36	1.59	2.27	1.46	6.50	1.32	6.03	1.95	5.97	1.73	6.67	.77	5.00	1.80
Tukey & Scheffe							(not available)									
Interaction 2-way	5.74	.95	5.23	1.40	4.83	1.32	6.66	.63	6.51	.94	5.57	1.70	6.74	.65	5.37	1.66
Chi-Square	6.44	.74	6.22	.93	5.74	1.22	5.58	2.09	6.68	.76	5.24	1.77	6.82	.57	5.35	1.26
Data Analysis	6.03	1.13	6.20	.75	4.93	1.84	6.10	1.90	6.10	1.56	5.33	1.99	6.57	1.20	5.47	1.50
Course Average	5.61		5.44		4.11		5.77		6.43		5.66		6.76		5.58	

USEFUL	"I found the material in this chapter was quite useful in helping me learn basic statistics."
INTEREST	*"The material in this chapter was quite boring."
EASY	"I found this chapter easy to understand."
NEW	"I previously had not been taught any of the material in this chapter."
HELP AVAIL.	*"There was never anyone available in the terminal room to help answer my questions about this chapter."
HELP ND.	"I feel that someone who is knowledgeable in the course content of this chapter should always be available in the terminal room."
HELP KNW.	*"Whenever there was a course content specialist available in the terminal room that person very often did not know the material very well."
CLASS-CAI	*"I feel that it would have been easier for me to learn the material in this chapter by classroom instruction rather than by computer-assisted instruction."

*These items were reflected for the data given above.

perceive the situation. The highest ratings of 6.36 for the Descriptive Statistics section seems incongruous, because this section contains only very basic arithmetic and statistical concepts. This high rating may be because this was the first topic of the course and for most students, their first use of CAI. It may also reflect the poor arithmetic competencies of some students.

General Observations

In both the 1975 and 1976 testing of the course the authors were surprised to find that the level of ego involvement of the students in the course was very high. Students appeared to set for themselves a very high level of performance on examinations, much higher than what the instructors had been accustomed to when using the lecture method.

Although no pre-measures were taken on either the 1975 or 1976 summer groups, the authors felt that the 1976 group of students had a larger proportion of students with poor backgrounds in basic arithmetic skills, as well as a higher level of anxiety towards their deficiencies. The authors suspected that some of these students had postponed taking the course until it was one of the last needed to complete their academic program. For these students progress in the first section of STAT1 was hampered by a lack of arithmetic skills generally considered to have been acquired at the grade seven or eight level. Here are some examples obtained by the instructors while working on an individual basis with such students:

- 1) -not being able to convert a fraction to a decimal
- 2) -not being able to place negative and positive integers into an ascending sequence. (To the question, "which is smallest -200, 0, or 200," the student replied 0!)
- 3) -not knowing that division by zero does not produce zero.
- 4) -unable to handle simple arithmetic involving fractions
- 5) -not realizing that division by 2 is the same as multiplying by $1/2$
- 6) -unable to multiply or divide using signed numbers
- 7) -unable to estimate the correct position of a decimal point in an arithmetic operation

Frequently students having these deficiencies would rationalize them by claiming that, when they finish their training, they would not be in a position which required such skills! Needless to say, the authors were dismayed with such an attitude.

Perhaps because of the limited amount of time available in a Summer Session, coupled with poor entry skills, the authors perceived some frustrations on the part of the poorer students. These frustrations manifested themselves in the following ways:

- 1) -more frequent claims that the computer was wrong, or that it was not treating them in a consistent and fair manner

- 2) -claims that a correct answer was given, but that the computer marked it incorrectly
- 3) -a student taking an inordinate amount of time to answer questions on an examination. (This problem became compounded because eventually the student would have to sign-off before the examination was completed. Upon re-entry into the examination at another time, the computer re-randomized the block of 15 questions in which the student was previously working. This led to claims that the computer was "now giving me different questions," and, that it was "not fair because I have to answer some questions twice!")
- 4) -impulsive behavior leading to the position that an incorrect answer was not intended and a request for additional marks on the examination.

Indeed, such inadequacies as illustrated above raised in the authors' minds the quality of the research that these students would eventually do to complete their thesis requirements, as well as why progress of research was so slow in the educational field.

The authors wish to note, however, that the majority of students, and particularly those from the Faculty of Physical Education, were a pleasure to work with.

Current Use of STAT1

STAT1 is being used again during the winter of 1976 by students in Educational Psychology 502, and by the Department of Elementary Education in a course entitled, Curriculum and Instruction 502. The latter Department is using only a portion of the course content. Some students from other departments are also using the course to upgrade their basic knowledge in the area. Because of the short interval of time between the 1976 Summer Session and classes beginning in September 1976, changes deemed desirable to the course on the basis of the summer testing could not be made. Such changes are being made by the authors, but will take most of the winter before they are completed.

APPENDIX A. STAT1 1975 Correlation Matrix

APPENDIX B. STAT1 1976 Correlation Matrix

STAT1 1975

CORRELATIONS AMONG EXAMINATIONS (N=32)

Test	1	2	3	4	5	6	7	8	9	10	11	12
1. Description	--	.21	.28	.34	.56	-.05	.40	.37	.42	.16	.53	.68
2. Prediction		--	.66	.55	.26	.36	.36	.13	.22	-.07	.39	.52
3. Correlation #1			--	.49	.38	.45	.38	.08	.30	.02	.48	.58
4. Correlation #2				--	.33	.13	.39	.30	.29	-.09	.43	.51
5. Hypothesis Testing					--	.11	.22	.36	.21	.06	.70	.71
6. Anova						--	.33	-.02	-.11	-.05	.30	.35
7. Scheffe							--	.52	.12	-.16	.38	.63
8. Interaction								--	.11	-.12	.38	.61
9. Chi-Square									--	-.16	.13	.33
10. Research Design										--	.09	.13
11. Final											--	.83
12. Final Mark												--

STAT1 SUMMER 1976

CORRELATIONS AMONG EXAMINATIONS

Test	1	2	3	4	5	6	7	8	9	10	11	12
1. Description	--	.43	.59	.67	.67	.34	.47	.45	.0	.38	.76	.78
2. Prediction		--	.37	.37	.29	.48	.48	.28	.01	-.04	.48	.54
3. Correlation #1			--	.64	.67	.21	.30	.54	.23	.26	.55	.69
4. Correaltion #2				--	.61	.29	.63	.53	.16	.39	.78	.84
5. Hypothesis					--	.23	.24	.48	.08	.35	.69	.72
6. Anova						--	.46	.52	.03	.23	.33	.56
7. Scheffe							--	.50	.26	.05	.53	.71
8. Interaction								--	.13	.41	.52	.77
9. χ^2									--	.07	-.01	.22
10. Research Design										--	.31	.47
11. Final Exam											--	.87
12. Final Mark												--